Comparative Analysis of Evolutionary Mechanisms of the Hemagglutinin and Three Internal Protein Genes of Influenza B Virus: Multiple Cocirculating Lineages and Frequent Reassortment of the NP, M, and NS Genes

STEPHEN E. LINDSTROM, YASUAKI HIROMOTO, HIDEKAZU NISHIMURA, TAKEHIKO SAITO,
REIKO NEROME, AND KUNIAKI NEROME*

Department of Virology I, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo 162-8640, Japan

Phylogenetic profiles of the genes coding for the hemagglutinin (HA) protein, nucleoprotein (NP), matrix (M) protein, and nonstructural (NS) proteins of influenza B viruses isolated from 1940 to 1998 were analyzed in a parallel manner in order to understand the evolutionary mechanisms of these viruses. Unlike human influenza A (H3N2) viruses, the evolutionary pathways of all four genes of recent influenza B viruses revealed similar patterns of genetic divergence into two major lineages. Although evolutionary rates of the HA, NP, M, and NS genes of influenza B viruses were estimated to be generally lower than those of human influenza A viruses, genes of influenza B viruses demonstrated complex phylogenetic patterns, indicating alternative mechanisms for generation of virus variability. Topologies of the evolutionary trees of each gene were determined to be quite distinct from one another, showing that these genes were evolving in an independent manner. Furthermore, variable topologies were apparently the result of frequent genetic exchange among cocirculating epidemic viruses. Evolutionary analysis done in the present study provided further evidence for cocirculation of multiple lineages as well as sequestering and reemergence of phylogenetic lineages of the internal genes. In addition, comparison of deduced amino acid sequences revealed a novel amino acid deletion in the HA1 domain of the HA protein of recent isolates from 1998 belonging to the B/Yamagata/16/88-like lineage. It thus became apparent that, despite lower evolutionary rates, influenza B viruses were able to generate genetic diversity among circulating viruses through a combination of evolutionary mechanisms involving cocirculating lineages and genetic reassortment by which new variants with distinct gene constellations emerged.

Influenza type A and B viruses share many characteristics both biologically and biochemically. For example, both viruses possess a segmented genome consisting of eight negative-strand RNA segments which encode three polymerase proteins associated with polymerase activities, a hemagglutinin glycoprotein (HA), a neuraminidase glycoprotein (NA), a nucleoprotein (NP), a matrix protein (M1), and two nonstructural proteins (NS1 and NS2). However, there are also significant differences between the epidemiology, evolutionary patterns, and host ranges of influenza A and influenza B viruses. In fact, influenza A viruses are subdivided into 15 HA and 9 NA subtypes, all of which have been isolated from aquatic avian species and many of which also infect other avian and mammalian species, including humans, horses, mink, whales, swine, and seals (11, 15, 22, 30, 36). The segmented genome of influenza viruses allows different influenza viruses to exchange gene segments during coinfection of a cell (reassortment). Through genetic reassortment, pandemic influenza viruses with novel hemagglutinin subtypes (antigenic shift) which are able to evade established immunity may suddenly appear in humans. Antigenic shift has occurred at least twice in the twentieth century, both times resulting in major pandemics causing high morbidity and mortality in humans around the world. In 1957,

the emergence of H2N2 (Asian) virus was the result of reassortment between avian H2N2 and human H1N1 (Spanish) virus. Again in 1968, antigenic shift occurred as a result of reassortment between H2N2 (Asian) virus and an avian H3 virus to create the human H3N2 (Hong Kong) virus. In contrast, influenza B viruses, which have been isolated only from humans, consist of a single HA and NA type and thus are incapable of antigenic shift. Furthermore, due to high evolutionary rates of influenza A viruses, the HA and NA proteins are able to evade established immunity in humans by gradually changing their antigenic profile (antigenic drift). By this mechanism, the HA gene of human influenza H3N2 viruses has evolved in essentially a sequential manner to generate new antigenic strains which displace variants of the previous season (8–10, 27). Although the HA protein of influenza B viruses undergoes antigenic drift, evolution of the HA protein of influenza B viruses has been characterized by a lower rate of antigenic change and cocirculation of antigenic variants for considerable periods of time (20, 39, 46). In particular, since the mid-1980s, the HA gene of influenza B viruses has been shown to have evolved into two distinct lineages represented by B/Yamagata/16/88- and B/Victoria/2/87-like viruses and to have demonstrated a mechanism of systematic insertion and deletion of nucleotides (20, 21, 32, 39). It was recently observed that outbreaks of influenza B virus in Japan tend to occur in association with H3N2 viruses, with influenza B virus activity peaking after that of H3N2 virus, usually in March (32). Nevertheless, it is still unclear how, despite slow evolutionary change and the lack of antigenic shift, influenza B viruses are

* Corresponding author. Mailing address: Department of Virology I, National Institute of Infectious Diseases, 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan. Phone: (03) 5285-1111. Fax: (03) 5285-1155. E-mail: knerome@nih.go.jp.

TABLE 1. Accession numbers of influenza B virus sequence data determined in this study and abbreviations of the virus isolates

Strain	Abbreviation	Sequence accession no.			
		Segment 4 HA gene	Segment 5 NP gene	Segment 7 M gene	Segment 8 NS gene
B/Bangkok/64	Ban64	AF101066			
B/Osaka/70	Osa70	AF101067			
B/Gifu/73	Gif73	AF101068			
B/Guma/73	Gum73	AF101069			
B/Kanagawa/3/76	Kan76	AF101070			
B/Norway/4/84	Nor84	AF101071	AF100357	AF100374	AF100393
B/Ibaraki/2/85	Iba85		AF100358	AF100375	AF100394
B/Victoria/2/87	Vic87		AF100359	AF100376	
B/Aichi/5/88	Aic88		AF100360	AF100377	AF100395
B/Yamagata/16/88	Yam88			AF100378	AF100396
B/Panama/45/90	Pan90			AF100379	AF100397
B/Mie/1/93	Mie93		AF100361	AF100380	AF100398
B/Guangdong/8/93	Gua93		AF100362	AF100382	AF100400
B/Beijing/184/93	Bei93		AF100367	AF100381	AF100399
B/Guangdong/5/94	Gua94		AF100363	AF100383	AF100401
B/Harbin/7/94	Har94		AF100364	AF100384	AF100402
B/Argentina/218/97	Arg97	AF100347	AF100365	AF100385	AF100403
B/Beijing/243/97	Bei97		AF100366	AF100386	AF100405
B/Henan/22/97	Hen97	AF100349	AF100368	AF100387	AF100404
B/Nara/4/97	Nar97	AF100354			
B/Hiroshima/97/97	Hir97	AF100356			
B/Shiga/44/98	Shi4498	AF100353			
B/Chiba/447/98	Chi98	AF100348	AF100369	AF100388	AF100408
B/Nagano/2038/98	Nag98	AF100350	AF100370	AF100389	AF100409
B/Shiga/51/98	Shi5198	AF100352	AF100371	AF100390	AF100406
B/Shiga/T30/98	Shi3098	AF100351	AF100372	AF100391	AF100407
B/Yamanashi/166/98	YN98	AF100355	AF100373	AF100392	AF100410

able to continue to cause seasonal epidemic episodes in humans.

In addition to differences in host range specificity and evolutionary patterns of the HA genes, influenza A and B viruses show further differences in their protein coding strategies. RNA segment 6 of influenza A virus is monocistronic, coding for the NA protein, while segment 6 of influenza B virus is bicistronic, possessing two overlapping open reading frames (ORFs) which code for the NA and NB proteins. The NB protein of influenza B viruses is a membrane protein which is believed to serve a function similar to that of the M2 protein of influenza A viruses (1, 2, 42, 44). Furthermore, although RNA segment 7 of both influenza A and B viruses codes for the M1 (matrix) protein, the organizations of their respective M2 genes are quite different. Processing of the M2 gene of influenza A viruses involves posttranscriptional splicing of the mRNA to obtain the M2 ORF. The M2 protein subsequently shares the first 14 amino acids (aa) with the M1 protein, while the remaining 88 aa are unique to the M2 protein (17, 24). On the other hand, RNA segment 7 of influenza B viruses is bicistronic, containing tandem cistrons characterized by overlapping of the termination codon of the M1 gene and the initiation codon of the BM2 gene, resulting in a rare coupled termination-initiation mechanism of viral gene expression (16). Although the NB protein of influenza B viruses is believed to form a membrane ion channel similar to that of the M2 protein of influenza A viruses (42, 44), influenza A viruses have no counterpart for the M2 protein of influenza B viruses (BM2), whose function remains unknown.

Although the phylogenetic patterns of the HA genes of influenza B viruses are well understood, evolutionary characteristics of the internal genes have not been well resolved. Recent analysis of seven influenza B virus NP genes provided evidence for multiple lineages and reassortment (19), while

Yamashita et al. (46) suggested multiple lineages of the NS gene. Recent advances in molecular techniques, such as rapid RNA purification, reverse transcription (RT)-PCR, direct sequencing of PCR amplicons, and automated nonradioisotopic DNA sequencing, have allowed for rapid and safe analysis of a large number of viral genes. In this report, we compare the phylogenetic profiles of the HA, NP, M, and NS genes of 20 influenza B viruses from 1940 to 1998 in a parallel fashion in order to better understand the evolutionary mechanisms of these viruses. A comparison of evolutionary mechanisms of influenza A and B viruses is also discussed.

MATERIALS AND METHODS

Viruses. The following influenza B virus strains whose genes were sequenced in this study were propagated in 11-day-old embryonated chicken eggs: B/Bangkok/64, B/Osaka/70, B/Gifu/73, B/Guma/73, B/Kanagawa/3/76, B/Norway/4/84, B/Ibaraki/2/85, B/Victoria/2/87, B/Aichi/5/88, B/Yamagata/16/88, B/Panama/45/90, B/Mie/1/93, B/Guangdong/8/93, B/Beijing/184/93, B/Guangdong/5/94, B/Harbin/7/94, B/Argentina/218/97, B/Beijing/243/97, B/Henan/22/97, B/Nara/4/97, B/Hiroshima/97/97, B/Shiga/44/98, B/Chiba/447/98, B/Nagano/2038/98, B/Shiga/51/98, B/Shiga/T30/98, and B/Yamanashi/166/98.

RNA extraction and nucleotide sequencing. Viral RNA was extracted from 100 μl of virus sample by use of a commercial kit (RNeasy; Qiagen) and suspended in a final volume of 40 μl of RNase-free distilled water. RT-PCR was performed by using a modified protocol of a commercial kit (RT-PCR kit with avian amyeloblastosis virus reverse transcriptase, version 2.1; Takara) as described previously (26). RT reactions were done with 20 pmol of a previously described universal influenza B virus RT primer (5′ AGCAGAAGC 3′) (47) and 9.5 μl of RNA sample/20 μl of RT reaction mixture. Aliquots of 5 μl of the resultant cDNA were then used in all subsequent PCR amplifications of overlapping cassettes covering the entire protein coding domain of the NP, M, and NS genes, as well as the HA1 domain of the HA gene, by using the following thermocycler program: initial denaturation at 94°C for 2 min followed by 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min and a final extension at 72°C for 10 min. Resultant amplicons were purified and sequenced directly by methods described previously (41) on an ABI377 or ABI310 autosequencer (Perkin-Elmer). Oligonucleotide primers for PCRs and sequencing reactions were designed to have annealing temperatures ranging from 62 to 66°C. Nucleotide sequences for all
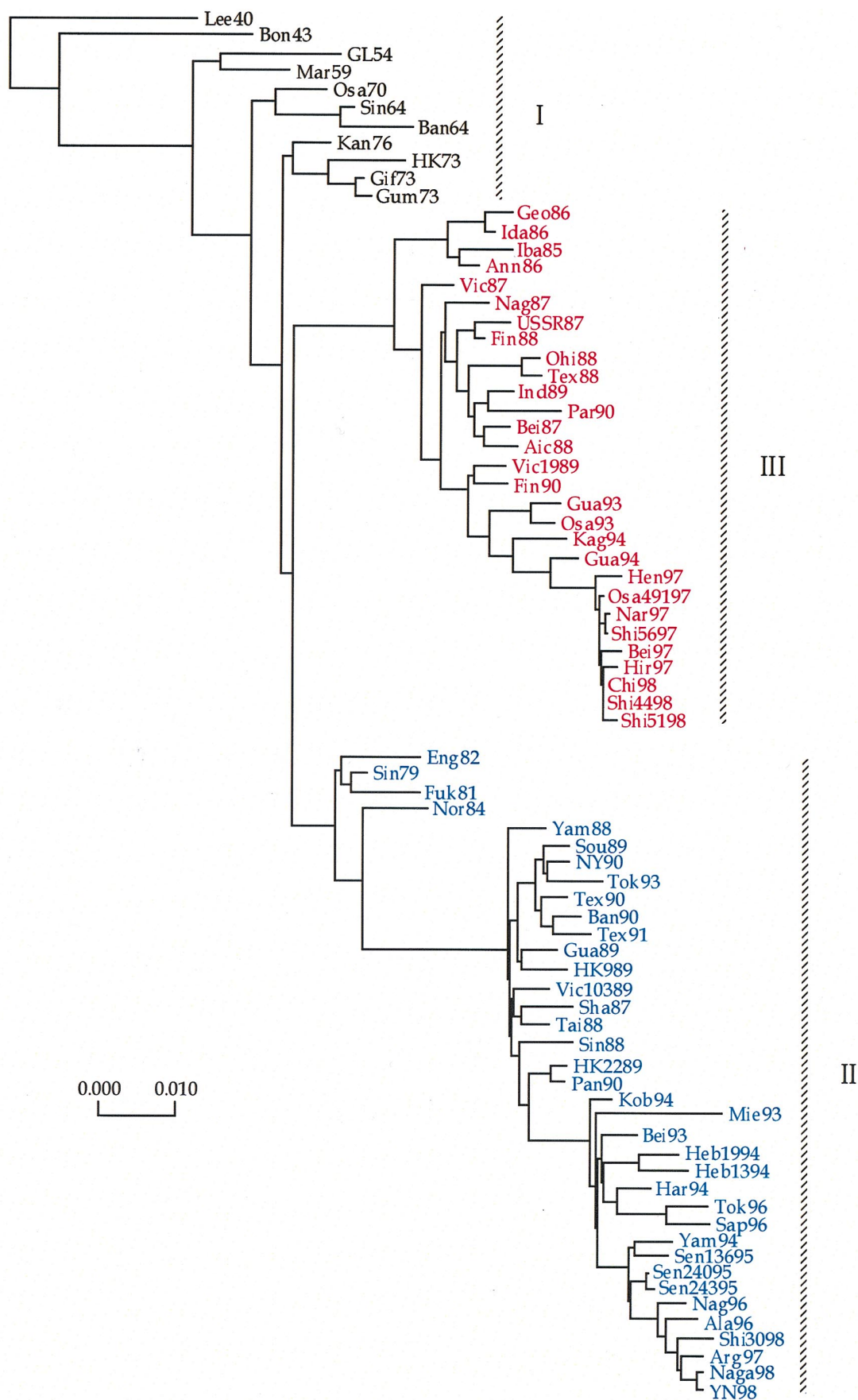
FIG. 1. Evolutionary tree based on the total number of nucleotide substitutions of the HA1 domain of the HA gene of human influenza B viruses constructed by neighbor-joining analysis. Isolates indicated in blue represent viruses belonging to lineage II, whereas those indicated in red represent viruses belonging to lineage III.

TABLE 2. Amino acid variation of HA protein (HA1 domain) of influenza B viruses[a]

Virus isolate	Lineage	HA protein amino acid at indicated position																													
		29	45	48	56	68	71	73	75	76	81	88	116	118	121	122	125	129	136	137	146	148	149	150	154	162	163	164	165	171	172
HK73	I	V	K	Q	N	G	K	M	T	I	A	K	N	R	A	R	T	T	I	V	V	N	G	N	A	K	~[b]	~	~	N	P
Iba85	III		K	K											T	H	I										N	N	N		
Vic87	III		K	K	A								H		T	H	I		K								N	D	N		
Aic88	III		K	K	A			T	N		V		H	G	T	H	I		K								N	N	N		S
Gua93	III		K	K				T	N		V	R	H		T	H	I	K	K	I	I						N	D	N		S
Gua94	III		K	K				T	N		V	R	H		T	H	I	K	K	I	I						N	D	N		S
Hen97	III		K	K				T	N		V	R	H		I	H	I	K	K	I							N	D	N		S
Bei97	III		K	K				T	N		V	R	H		I	H	I	K	K	I					V		N	D	N		S
Chi98	III		K	K				T	N		V	R	H		I	H	I	K	K	I							N	D	N		S
Shi5198	III		K	K				T	N		V	R	H		I	H	I	K	K	I							N	D	N		S
Sin79	II			K											T		I	R		I							~	~	D		
Nor84	II			K								R			T		I	R	R	L							~	~	~	D	
Yam88	II			K		A	M					R			T	H	I	R	R	L		S	R			R	~	~	D		
Pan90	II			K		A	M			V	T	R			T	Q	I	R	R	L		S	R	D		R	~	~	D		
Bei93	II			K			M			V	T	R			T	Q	I	K	R	L	A	S	R	S		R	~	D	N		
Mie93	II		N	K		A	M			V	T	R			T	Q	I	K	R	L	A	S	R	S		R	~	D	N	I	
Har94	II			K			M			V	T	R			T	Q	I	K	R	L	A	S	R	S		R	~	D	N		
Shi3098	II	A		K	T		M	V		I	T	R	K		T	Q	I	K	R	L	A	S	R	S			~	~	D		N
Arg97	II			K	T		M	V		I	T	R	K		T	Q	I	K	R	L	A	S	R	S			~	~	D		N
Nag98	II			K	T		M	V		V	T	R	K		T	Q	I	K	R	L	A	S	R	S		R	~	D	N		
YN98	II			K	T		M	V		V	T	R	K		T	Q	I	K	R	L	A	S	R	S			~	~	D		N

[a] Deduced amino acid differences in the HA1 domain of the HA protein of influenza B viruses. Only locations where differences from the sequence of HK73 were observed are indicated.

[b] ~, amino acid deletion.

oligonucleotide primers used for PCR amplification and sequencing are available from the authors upon request.

Phylogenetic analyses and evolutionary rates. Phylogenetic trees were constructed by using the neighbor-joining method (12, 31, 40) and bootstrap analysis (n = 500) to determine the best-fitting tree for each gene. Nucleotide distance matrices were estimated by the six-parameter-method (12) based on the number of total nucleotide substitutions, and evolutionary trees for the HA, NP, M, and NS genes were constructed (see Fig. 1 to 4, respectively) (13, 40). In addition to the sequence data determined in this study, previously reported nucleotide sequences for the HA gene (20, 21, 23, 32, 38, 39, 45, 46), NP gene (5, 7, 19, 28, 35, 37), M gene (4, 7, 14), and NS gene (3, 33, 46) were also used in the construction of phylogenetic trees. The abbreviations used for influenza B virus strains are as follows: B/Lee/40 (Lee40), B/Bon/43 (Bon43), B/Great Lakes/54 (GL54), B/Maryland/59 (Mar59), B/Singapore/64 (Sin64), B/Ann Arbor/66 (Ann66), B/Russia/69 (Rus69), B/Hong Kong/8/73 (HK73), B/Yamagata/73 (Yam73), B/Baylor/4/78 (Bay78), B/Singapore/222/79 (Sin79), B/Paris/1/79 (Par79), B/Fukuoka/80/81 (Fuk81), B/England/222/82 (Eng82), B/Houston/1851/84 (Hou84), B/Georgia/1/86 (Geo86), B/Idaho/1/86 (Ida86), B/Ann Arbor/1/86 (Ann86), B/Nagasaki/1/87 (Nag87), B/Beijing/1/87 (Bei87), B/USSR/2/87 (USSR87), B/Shanghai/2/87 (Sha87), B/Finland/56/88 (Fin88), B/Taiwan/7/88 (Tai88), B/Singapore/7/88 (Sin88), B/India/3/89 (Ind89), B/Victoria/19/89 (Vic1989), B/Victoria/103/89 (Vic10389), B/South Dakota/5/89 (Sou89), B/Guangdong/55/89 (Gua89), B/Paris/329/90 (Par90), B/Finland/151/90 (Fin90), B/New York/3/90 (NY90), B/Texas/4/90 (Tex90), B/Bangkok/163/90 (Ban90), B/Hong Kong/22/89 (HK2289), B/Hong Kong/9/89 (HK989), B/Texas/1/91 (Tex91), B/Osaka/1/93 (Osa93), B/Tokyo/101/93 (Tok93), B/Mie/1/93 (Mie93), B/Kagoshima/15/94 (Kag94), B/Kobe/1/94 (Kob94), B/Hebei/19/94 (Heb1994), B/Hebei/3/94 (Heb394), B/Harbin/7/94 (Har94), B/Yamagata/74/94 (Yam94), B/Sendai/240/95 (Sen24095), B/Sendai/243/95 (Sen24395), B/Sendai/136/95 (Sen13695), B/Tokyo/942/96 (Tok96), B/Sapporo/1/96 (Sap96), B/Nagasaki/1/96 (Nag96), B/Alaska/12/96 (Ala96), B/Osaka/491/97 (Osa97). The remaining strain abbreviations are indicated in Table 1.

Evolutionary rates based on the total number of nucleotide substitutions in the HA, NP, M, and NS genes as well as the number of amino acid substitutions in the HA, NP, M1, BM2, NS1, and NS2 proteins were calculated by plotting the evolutionary distance (number of substitutions per site) of each virus from a putative origin against the year of isolation and determining the slope of the best-fit line by regression analysis.

Nucleotide sequence accession numbers. Nucleotide sequence data determined in this study will appear in the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases under the accession numbers listed in Table 1.

RESULTS

Analysis of the HA gene. Analysis of the HA gene of 77 influenza B viruses (Fig. 1) was undertaken in order to provide a basis for comparison of evolutionary patterns of the internal genes, as well as to understand the phylogenetic location of the HA genes of recent isolates from 1997 to 1998. In support of previous analyses (20, 21, 32, 38, 39), the HA gene was demonstrated to have evolved into three major lineages. Lineage I included earliest isolates from 1940 until the mid-1970s, while more recent isolates were found to form two distinct lineages which appeared to have diverged prior to the isolation of Sin79. These results support previous reports describing cocirculation of antigenically and phylogenetically distinct Yamagata/16/88-like (lineage II) and Victoria/2/87-like (lineage III) variants (20, 21, 32, 38, 39). Phylogenetic divergence of lineages II and III was supported by calculated bootstrap probabilities of 95 and 100%, respectively. Analysis of recent influenza B virus isolates from 1997 to 1998 revealed that variants of both lineage II and lineage III were isolated in Japan in 1998, forming new clades at the ends of each lineage. Also, viruses of both lineages were isolated from the same geographic region; Shi4498 and Shi5198 belonged to lineage III, while a third virus isolated from the same area, Shi3098, was included in lineage II.

A comparison of predicted amino acid sequences of the HA1 domain of the HA proteins of recent isolates to that of HK73 (Table 2) revealed many conserved changes which were characteristic of each lineage. However, it was interesting to observe that as many as 12 aa changes occurred independently at the same positions in viruses of both lineages. For example, even though phylogenetically distinct viruses Shi5198 (lineage III) and Shi3098 (lineage II) differed by 32 aa (9.2%), 11 of these changes were at similar locations. Independent changes at similar sites may be indicative of lower functional constraints on amino acids at these positions or, possibly, escape from immune pressure directed to these sites. However, the latter argument does not explain why viruses of both lineages demonstrated similar amino acid changes at positions 88 (K to R [K-R]), 129 (T-K, lineage III; R-K, lineage II), and 267 (V-I). The HA protein of influenza B viruses is characterized by the insertion and deletion of amino acids at positions 163 to 165

TABLE 2—*Continued*

HA protein amino acid at indicated position

175	179	180	183	184	185	190	197	198	199	200	202	203	209	220	230	233	235	238	240	250	255	262	267	274	276	287	293	298	302	313	330
V	Y	I	K	G	E	V	D	E	T	Q	V	K	K	V	G	N	A	E	L	Y	P	A	V	W	A	L	A	E	G	E	K
	F	E															G				H	S	T	I		S	P	F		A	A
		E					S										G					S	T	I							
		E					S										G					S	T	I			R				
I		E					N				A						G					S	T	I							N
I		E					N		I		A					T	G					S	T	I							
I		E	I				N		N		A					T	G					S	T	I							
I		E					T				A					T	G					S	T	I							
I		E					N				A					T	G					S	T	I							
I		E					N		A		A					T	G					S	T	I							
							T									T	G					V									
I	C						N	K	A							T	G					V								A	
								K			K		N		D	T	G					V							A		
							N	K			K	N	N			T	G					V							A		
							N	K	I	R	H	T	N	N	D	T	G	R				V									
		V		E		G	N	K	A		K	N	N		D	T	G					V									
	H			E			N	K			K	N	N	I	D	T	G					V	I								
	H			E			S	K			K	N	N		D	T	G					V	I								
	H			E			N	K			K	N	N		D	T	G					V	I								
	H			E			N	K			K	N	N		D	T	G					V	I								

(20, 32, 39). It was, therefore, relevant that three of the four most recent viruses of lineage II, Arg97, Shi3098, and YN98, were found to have a novel deletion at position 162.

Estimation of nucleotide and amino acid substitution rates of the HA1 domain of the HA gene of influenza B viruses (Table 3) revealed that the evolutionary rates of both lineages calculated here were somewhat lower than those previously reported. Nucleotide substitution rates for lineages II and III were estimated to be 2.41×10^{-3} nucleotide substitutions/site/year (ns/st/yr) and 1.39×10^{-3} ns/st/yr, respectively, which are the converse of previously described rates of 3.42×10^{-3} and 4.17×10^{-3} ns/st/yr, respectively (39). These differences were likely attributable to inclusion of more sequence data collected over a longer period of time in the calculation of this report. It was also apparent that the HA1 domain of the HA gene of influenza B viruses was evolving at a rate about three times lower than that of influenza A H3N2 viruses (5.7×10^{-3} ns/st/yr) (9). At the amino acid level, the HA protein of lineage II was observed to be evolving at a rate of 3.36×10^{-3} amino acid substitutions/site/year (aas/st/yr), which was slightly higher than that of lineage III (2.18×10^{-3} aas/st/yr) and was consistent with the amino acid variability observed in this study and previously (38).

Analysis of the NP gene. Phylogenetic analysis of the NP gene of 24 influenza B viruses isolated since 1940 revealed evolutionary profiles similar to those of the HA gene, showing that the NP gene has evolved into three major lineages (Fig. 2). Lineage I was represented by the classical isolates Lee40 and Ann66, while more recent viruses were observed to have diverged into two distinct lineages (lineages II and III). However, unlike the evolutionary patterns of the HA gene, NP genes of lineage III appeared to have circulated for a relatively short period of time, since this lineage consisted entirely of viruses isolated between 1984 and 1988. Nevertheless, lineages II and III were demonstrated to have bootstrap confidence levels of 99 and 100%, respectively. Also, as can be seen in Fig. 2, the constituent viruses of each lineage contrasted considerably with those of the HA gene. For example, the NP genes of viruses of HA lineage III and HA lineage II were not restricted to either NP lineage. Rather, NP lineages II and III consisted of a mixture of viruses from both HA lineage II and HA lineage III, reflecting frequent reassortment of these genes among influenza B viruses. These results confirm evidence for reassortment among influenza B viruses based on the evolutionary position of the NP gene of Tex88 observed by Jambrina et al. (19). Further analysis of the NP gene in this study demonstrated that reassortment appeared to have occurred quite regularly among influenza B viruses, including the most recent isolates from 1998. Interestingly, all NP genes of recent viruses belonged to lineage II, forming two branch clusters, IIi and IIii, representing viruses of HA lineage II and HA lineage III, respectively.

Alignment of deduced amino acid sequences of the NP proteins of both lineages with that of Ann66 (Table 4) revealed lower variability in the NP protein when compared to the HA protein. There were a total of 12 conserved (i.e., shared by three or more viruses) aa differences (2.1%) between the NP proteins of lineage II and lineage III, while the NP proteins of recent isolates of branch clusters IIi and IIii differed by only 4 conserved aa (0.7%) at positions 17 (T-A, lineage IIii), 171 (I-V, lineage IIii), 233 (L-I, lineage IIi) and 520 (K-R, lineage IIi).

The evolutionary rate of the NP gene of influenza B viruses (Table 3) of lineage II (0.95×10^{-3} ns/st/yr) was approximately one-fourth that previously estimated (3.6×10^{-3} ns/st/yr) (37). Also, this rate was about one-half that of the NP gene of influenza A viruses (2.3×10^{-3} ns/st/yr), while the NP protein of influenza B viruses (0.26×10^{-3} aas/st/yr) was calculated to be evolving at a rate which was almost 1/10 that of influenza A viruses (2.1×10^{-3} aas/st/yr) (43).

Analysis of the M gene. Construction of the evolutionary tree of the M gene of influenza B viruses included the complete protein coding region of 1,076 bp comprising the M1 and M2 ORFs of the M gene of 24 viruses (Fig. 3). The M gene was found to have evolved into three distinct lineages which, with the exception of one virus, were topologically identical to those

TABLE 3. Evolutionary rates of influenza B virus genes[a]

Gene	Evolutionary lineage	Nucleotide substitution rate[b] (nt/st/yr [10^{-3}])	Amino acid substitution rate (aat/st/yr [10^{-3}])
HA	II	2.41	3.36
	III	1.39	2.18
NP	II	0.95	0.26
	III	ND[c]	ND
M (complete)	II	1.09	
	III	1.31	
M1	II	0.75	ND
	III	0.30	ND
M2	II	2.01	3.80
	III	4.14	3.46
NS (complete)	III	0.87	
	IV	0.45	
NS1	III	1.03	1.34
	IV	0.37	0.49
NS2	III	0.30	0.76
	IV	0.29	0.52

[a] Nucleotide and amino acid substitution rates of lineages II and III of the HA, NP, and M genes as well as lineages III and IV of the NS gene were estimated by determining the evolutionary distances of each virus from the putative node of divergence.

[b] Nucleotide evolutionary rates for the HA gene were based on the entire region of the HA1 domain. Rates for all other genes were based on the entire protein coding regions.

[c] ND, evolutionary rates could not be estimated because of a lack of correlation between evolutionary distances and time between isolation of strains. The M1 protein showed essentially no evolution since there were no conserved amino acid differences in isolates since 1966.

of the NP gene, with lineages II and III having bootstrap confidence values of 96 and 100%, respectively. Similar to the NP gene, the M genes of most recent viruses were all located in lineage II, forming two branch clusters, IIi and IIIi, whose HA genes belonged to HA lineages II and III, respectively. However, it was interesting to note that the M gene of Gua93 was phylogenetically located in lineage III, indicating a re-emergence of the M gene of lineage III after a 9-year hiatus. This was in contrast to the NP gene of this virus, which was included in lineage II of the evolutionary tree of the NP gene.

A comparison of predicted amino acid sequences of the M proteins of influenza B viruses (Table 5) revealed almost complete conservation of the M1 proteins among influenza B viruses when compared with that of Ann66. Regardless of phylogenetic distinctions, there were no conserved amino acid differences found in the M1 proteins of viruses isolated during a 32-year period from 1966 to 1998. Similar to the conserved nature of the M1 protein of influenza A viruses (18, 26), the lack of variability in the M1 protein of influenza B viruses suggested high functional constraints on this protein. Although the function of the BM2 protein is not yet well understood, all viruses contained a conserved BM2 ORF of 330 nucleotides coding for a predicted polypeptide of 109 aa, supporting the BM2 ORF described by Horvath et al. (16). In contrast to the M1 protein, high amino acid variability was demonstrated among BM2 proteins. Eleven conserved amino acid differences (10.1%) were observed between lineages II and III, while most recent isolates varied by five conserved amino acids (4.6%).

TABLE 4. Amino acid variation of the NP protein of influenza B viruses[a]

Virus isolate	Lineage	9	11	14	17	21	22	26	28	34	56	60	61	63	65	66	129	131	158	171	182	233	239	244	252	272	322	328	339	374	397	406	454	499	513	515	517	520	531	534	535	537	544
Ann66		I	T	I	T	I	T	S	A	K	T	A	I	G	R	T	Y	K	V	I	Y	L	A	I	A	R	V	V	N	M	Q	T	G	N	N	N	F	K	I	V	D	P	N
Nor84	III		A		T					R			D			F	F						T	T		K			Y		L								T	I	E	Q	
Iba85	III		A		T					R			D			F	F						T	T		K			Y		L					G			T	I	E		S
Vic87	III		A	A						R			D			F	F						T	T		K			Y		L								T	I	E		
Aic88	III		A	A	T					R			D			F	F						T	T		K			Y		L								T	I	E		
Sin79	II		A			I				R			D		K	F	F									K			Y		L								T		E		
Yam88	II								T	R			D		K	F	F	R						V		K		L	Y		L	I							T		E		
Pan90	II								T	R			D		K	F	F								V	K			Y		L								T		E		
Gua93	II			T				G	T	R			D		K	F	F									K			Y		L								T		E		
Bei93	II			T			T		T	R			D		K	F	F							V		K			Y		L								T		E		
Mie93	II								T	R			D		K	F	F		I							K			Y		L								T		E		
Gua94	II		A						T	R			D		A	F	A									K	I		Y		L								T		E		
Har94	II								T	R			D		K	F	F					I				K	I		Y		L				S				T		E		
Shi3008	IIi	V							T	R			D		K	F	F					I				K			Y		L		S			S			T		E		
Arg97	IIi				T				T	R			D		K	F	F					I				K			Y		L		S	S			L	R	T		E		
Nag98	IIi								T	R			D		K	F	F					I				K			Y		L			S			L	R	T		E		
YN98	IIi								T	R			D		K	F	F					I				K			Y		L			S				R	T		E		
Hen97	IIIi		A	A					T	R			D		K	F	F			V						K			Y		L								T		E		
Bei97	IIIi		A	A					T	R			D	E	K	F	F			V	C					K			Y		L								T		E		
Chi98	IIIi		A	A					T	R			E	E	K	F	F			V						K			Y		L								T		E		
Shi5198	IIIi		A	A					T	R			D		K	F	F			V						K			Y	V	L								T		E		

[a] Deduced amino acid differences in the NP proteins of influenza B viruses. Only locations where differences from the sequence of Ann66 were observed are indicated.
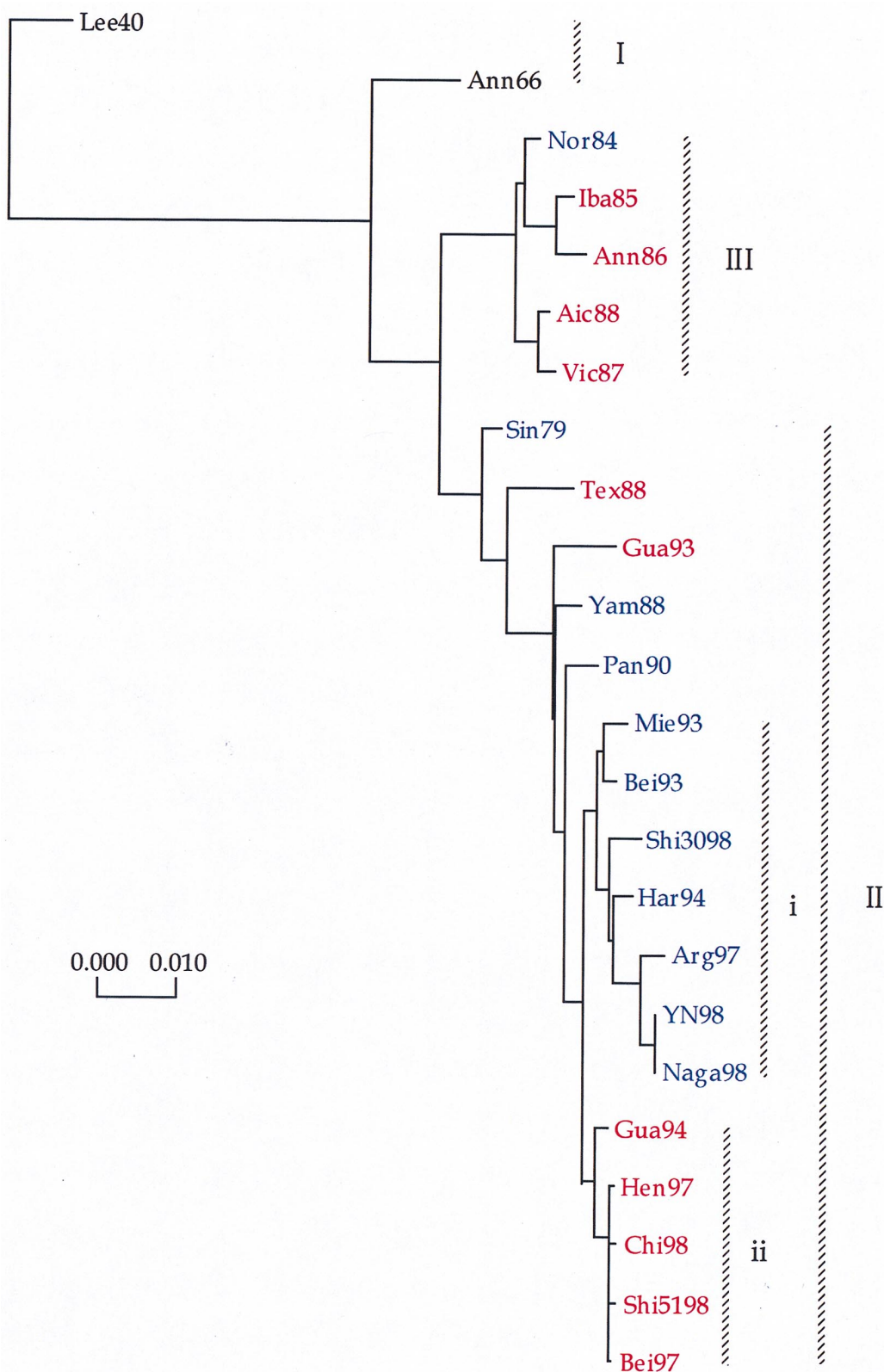
FIG. 2. Evolutionary tree of the NP gene of human influenza B viruses based on the total number of nucleotide substitutions in the complete protein coding region constructed by neighbor-joining analysis. Isolates indicated in red represent viruses whose HA gene belonged to lineage I, while isolates indicated in blue represent those whose HA gene belonged to lineage II.
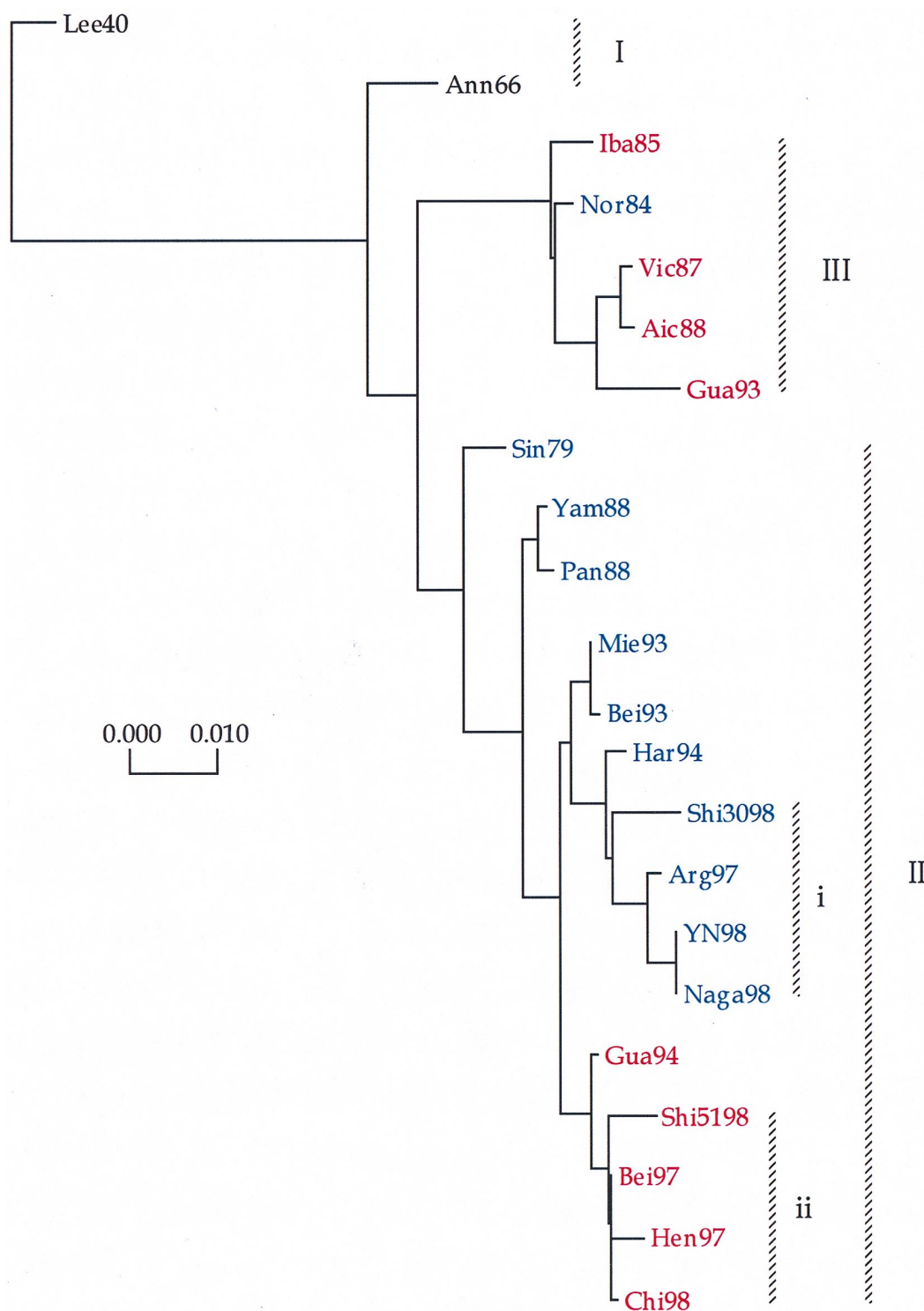
FIG. 3. Evolutionary tree of the M gene of human influenza B viruses based on the total number of nucleotide substitutions in the complete protein coding region constructed by neighbor-joining analysis. Isolates indicated in red represent viruses whose HA gene belonged to lineage I, while isolates indicated in blue represent those whose HA gene belonged to lineage II.

Lineages II and III of the M gene of influenza B viruses demonstrated similar nucleotide substitution rates of 1.09×10^{-3} and 1.31×10^{-3} ns/st/yr (Table 3), respectively, which were comparable to that calculated for human A viruses (1.08×10^{-3} ns/st/yr) (18). Nucleotide sequences of the M1 ORF of B viruses of lineage II displayed conservative rates of 0.75×10^{-3} (lineage II) and 0.30×10^{-3} (lineage III) ns/st/yr, while evolutionary rates of the BM2 ORF of lineages II and III (2.01×10^{-3} and 4.14×10^{-3} ns/st/yr, respectively), on the other hand, were markedly higher. In addition, the rates of

TABLE 5. Amino acid variation of the M1 and BM2 proteins of influenza B viruses[a]

Virus isolate	Lineage	M1 protein		M2 protein																						
		97	135	11	21	24	29	34	36	37	41	42	59	73	76	77	78	79	80	82	86	100	105	107	109
Ann66	I	I	P	C	M	T	N	G	N	L	R	N	S	V	D	N	M	E	I	S	V	E	V	E	Q
Nor84	III						S								N										
Iba85	III						S								N										
Vic87	III						S							A	N	K									
Aic88	III						S							A	N	K									
Gua93	III				I	I	S							T	N										
Sin79	III		R																						H
Yam88	II									M	K								V						H
Pan90	II									M	K								V						H
Bei93	II									M	K	S							V						H
Mie93	II									M	K	S							V						H
Gua94	II						S	R		M	K	G	N						V						H
Har94	II									M	K	S							V						H
Shi3098	IIi			R						M	K	S							V		I		I		H
Arg97	IIi									M	K	S							V				I		H
Nag98	IIi	T								M	K	S							V	N			I		H
YN98	IIi	T								M	K	S							V	N			I		H
Hen97	IIIi						S		S	M	K	G	N					K	V			K	I		H
Bei97	IIIi						S			M	K	G	N						V			K			H
Chi98	IIIi						S			M	K	G	N						V			K			H
Shi198	IIIi						S			M	K	G	N				I		V			K		K	H

[a] Deduced amino acid differences in the M1 and M2 proteins of influenza B viruses. Only locations where differences from the sequence of Ann66 were observed are indicated.

change of the BM2 ORF of influenza B viruses were higher than that of the M2 coding region of influenza A viruses (1.36×10^{-3} ns/st/yr). At the protein level, the M1 protein of influenza B viruses displayed virtually no evolution, since there were no conserved amino acid changes observed among virus isolates from 1966 to 1998. In contrast, the BM2 protein of lineages II and III was observed to be evolving at rates of 3.80×10^{-3} and 3.46×10^{-3} aas/st/yr, respectively, which were unexpectedly higher than those of the HA protein. Furthermore, these rates were about 2.5 times higher than that of the M2 protein of human A viruses (1.38×10^{-3} aas/st/yr) (18), although it should be noted that the M2 protein of influenza A viruses and the BM2 protein of influenza B viruses are not believed to share functionality.

Analysis of the NS gene. Calculation of the phylogenetic tree of the NS genes of influenza B viruses included 32 sequences determined in this report and previously (3, 33, 46) and revealed divergence into four major lineages (Fig. 4). Lineage I included the NS genes of the earliest isolates, Lee40 and GL54, with those of later viruses of the 1950s and 1960s being located in lineage II. Similar to the evolution of the NP and M genes, the NS gene of more recent viruses was characterized by phylogenetic divergence into two major lineages, supporting analysis by Yamashita et al. (46), who described multiple lineages of the NS gene based on the location of the NS gene of Hou84. Although the NS genes of recent isolates were found to be divided into two major lineages, evolutionary distances of NS genes of each lineage were often not proportional to their years of isolation, suggesting cocirculation of minor sublineages. Moreover, analysis of this report further revealed that viruses containing HA genes of HA lineages II and III were distributed throughout both NS gene lineages III and IV and that the general topology of the NS tree was quite distinct from those of the NP and M genes. Thus, the NS gene of influenza B viruses demonstrated a pattern of genetic reassortment

which was distinct from those of the NP or M genes. In particular, the NS gene of the representative strain of HA lineage III, Vic87, was located in NS lineage III together with that of Yam88, the representative strain of HA lineage II. Also, in contrast to the NP and M genes of isolates from 1997 and 1998, the NS genes of all isolates from 1997 and 1998 were located in lineage IV, forming minor clades, IVi and IVii, representing viruses of HA lineages III and II, respectively. Regardless, lineages III and IV appeared to have diverged prior to the isolation of HK73 with respective bootstrap probability values of 85 and 99%.

A considerable number of conserved amino acid changes were observed between the NS1 proteins of each lineage (Table 6). A total of 20 aa differences (7.1%) were observed between the NS1 protein of lineage III and lineage IV, while virus isolates from 1997 and 1998 of branch clusters IVi and IVii differed by 3 aa (1.1%). A fourth amino acid substitution was observed at position 252 (S-P) in two of four virus isolates of branch cluster IVi (Nag98 and YN98) and three of four virus isolates of branch cluster IVii (Bei97, Shi5198, and Chi98). Amino acid alignment of the NS2 proteins revealed a conserved sequence with two conserved changes (1.6%) between lineages III and IV and only 1 aa (0.8%) difference between that of viruses from 1997 and 1998 of sublineages IVi and IVii.

Based on the evolutionary distances of the NS genes, the nucleotide substitution rates for the NS genes of influenza B viruses of lineages III and IV were estimated to be 0.87×10^{-3} and 0.45×10^{-3} ns/st/yr, respectively. Although the NS gene of lineage IV appeared to be changing at a rate approximately two times higher than that of lineage III, both lineages were evolving at a noticeably lower rate than that of influenza A viruses (1.94×10^{-3} ns/st/yr) (29). Similar variability was subsequently observed between substitution rates of the NS1 gene-coding domain of lineages III (1.03×10^{-3} ns/st/yr) and IV (0.37×10^{-3} ns/st/yr). The NS2 coding region demon-
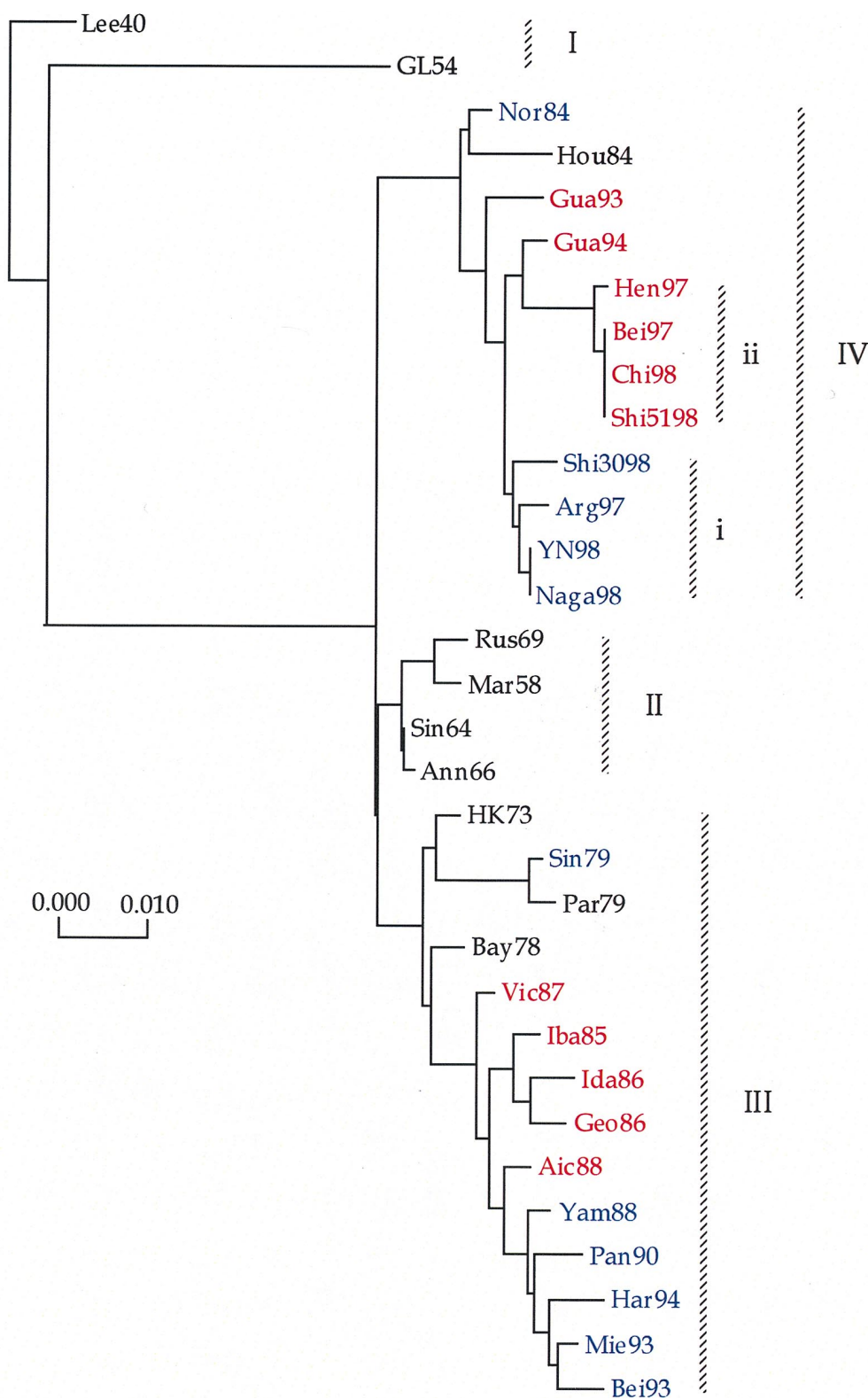
FIG. 4. Evolutionary tree of the NS gene of human influenza B viruses based on the total number of nucleotide substitutions in the complete protein coding region constructed by neighbor-joining analysis. Isolates indicated in red represent viruses whose HA gene belonged to lineage I, while isolates indicated in blue represent those whose HA gene belonged to lineage II.

strated the lowest rate of change since both lineages were estimated to be evolving at a rate of approximately 0.30×10^{-3} ns/st/yr. Similarly, amino acid substitution rates of the NS1 proteins of lineage III and IV were quite different from one another but markedly lower than that of human H3N2 viruses (3.6×10^{-3} aas/st/yr) (25). The NS2 protein of influenza B viruses revealed conservative substitution rates of 0.76×10^{-3} (lineage III) and 0.52×10^{-3} (lineage IV) aas/st/yr which were comparable to that of the NS2 protein of human H3N2 viruses (0.50×10^{-3} aas/st/yr) (25).

DISCUSSION

Comparison of the evolutionary profiles of the HA, NP, M, and NS genes of influenza B viruses revealed that these genes of recent isolates consistently divided into two distinct major lineages which were evolving independently. Also, analysis of recent influenza B viruses showed that viruses containing HA genes of both lineages cocirculated in Japan in 1998. It was also significant to reveal that 1998 viruses of lineage II demonstrated a novel amino acid deletion at position 162 of the HA1 domain of the HA protein. This is yet another example of the characterized systematic insertion and deletion of nucleotides in the protein coding region of the HA protein of influenza B viruses which potentially alters the antigenicity of these viruses (20, 32, 39).

In spite of similar patterns of genetic divergence, the respective evolutionary locations of the HA, NP, M, and NS genes of recent influenza B viruses differed considerably. Figure 5 illustrates the proposed evolution of influenza B viruses based on the phylogenetic characterization of the HA, NP, M, and NS genes of viruses examined in this study. Following phylogenetic divergence into two major lineages, genetic reassortment apparently occurred among cocirculating viruses of each lineage, giving rise to distinct viruses with novel genome constellations. It was apparent that variable genetic constellations observed among influenza B viruses were a reflection of frequent genetic reassortment among cocirculating viruses. In fact, only six virus isolates (Sin79, Yam88, Pan90, Mie93, Bei93, and Har94) contained genes which were consistently located in the same respective lineage, while the remaining virus isolates demonstrated reassortment of at least one gene segment. It was interesting to observe that divergent NP, M, and NS genes, represented by those of Nor84, were demonstrated to have been isolated independently of the earliest divergent HA gene of lineage III, represented by that of Iba85. Gene constellations of recent virus isolates from 1997 and 1998 provided additional evidence of genetic exchange among cocirculating viruses. While the HA genes of isolates from 1997 and 1998 belonged to two distinct lineages with high amino acid variability, the internal NP, M, and NS genes of these viruses, on the other hand, were less heterogeneous as they were included in similar lineages. Recently, evidence for reassortment of genes coding for the internal proteins of cocirculating human A H3N2 viruses was reported (26), although these genes were demonstrated to evolve essentially in a single major lineage. Observed reassortment among variable cocirculating internal genes of influenza B viruses was rather more similar to characterized reassortment among genetically variable influenza C viruses (6, 34). However, in contrast to that of influenza A and C viruses, reassortment among viruses of distinct lineages of influenza B viruses resulted in more pronounced protein variability.

It was further revealed that genes of a particular genetic lineage of the internal genes may be sequestered for a period of time and, through genetic reassortment, reemerge to circu-

TABLE 6. Amino acid variation of the NS1 and NS2 proteins of influenza B viruses[a]

Virus isolate	Lineage	NS1 protein	NS2 protein

Virus isolate	Lineage	2	3	36	69	92	97	102	104	106	107	112	113	115	116	119	120	121	123	126	127	128	131	136	138	139	140	141	142	167	179	182	184	202	207	211	215	217	247	252	279	2	3	18	35	41	59	80	87	98
Ann66	I	A	D	S	N	I	Y	M	N	S	N	N	C	N	C	T	D	Y	P	G	K	C	D	P	N	V	D	D	T	K	I	T	L	S	H	A	S	R	F	S	G	A	D	A	S	K	R	N	V	I
Nor84	IV			W	G																																									K				S
Gua93	IV			W	V						S									R					G																				K				
Gua94	IV			W					D		S									R					G																				K				
Shi3098	IVi				N	W					S	V								R					G																				K				
Arg97	IVi			W					D		S		Y							R					G															N			V		K				
Nag98	IVi			W				N	D		S		Y							R					G						M														K				
YN98	IVi			W		D				S			Y							R					G																				K				
YN97	IVi			W		D				S			Y							R					G																				K				
Hen97	IVii			W		D		S		S			Y	K			S		R					T											L										K				
Bei97	IVii			W		D		S		S			Y	K			S		R					T											L										K				
Shi5198	IVii			W		D		S		S			Y	K			S		R					T											L		N								K				
Chi98	IVii		D	W		D		S		S			Y	K			S		R					D											L										K				V
Sin79	II	E	W					S		F	Y		A				D	A				K	T	V												E								K				
Vic87	II		W											A								K	T																						K				
Iba85	II		W			V	D			P			A		A							Y																							K				
Aic88	II		W			V				P			A		A			G	A																										K				
Yam88	II		W									D			A							Y	G																						K				
Pan90	II		W			V				P			D		A				P	Y			D			G																			K				
Har94	II	E	W			V				P			Y	D				P	Y			N	D			G		H									N	K		L		R	E		K			A	V
Mie93	II	E	W			V				P			Y	D				P	Y				D			G																R	E		K				V
Bei93	II	E	W			V				P			D		A				P	Y			D			G																R	E		K			A	V

[a] Deduced amino acid differences in the NS1 and NS2 proteins of influenza B viruses. Only locations where differences from the sequence of Ann66 were observed are indicated.
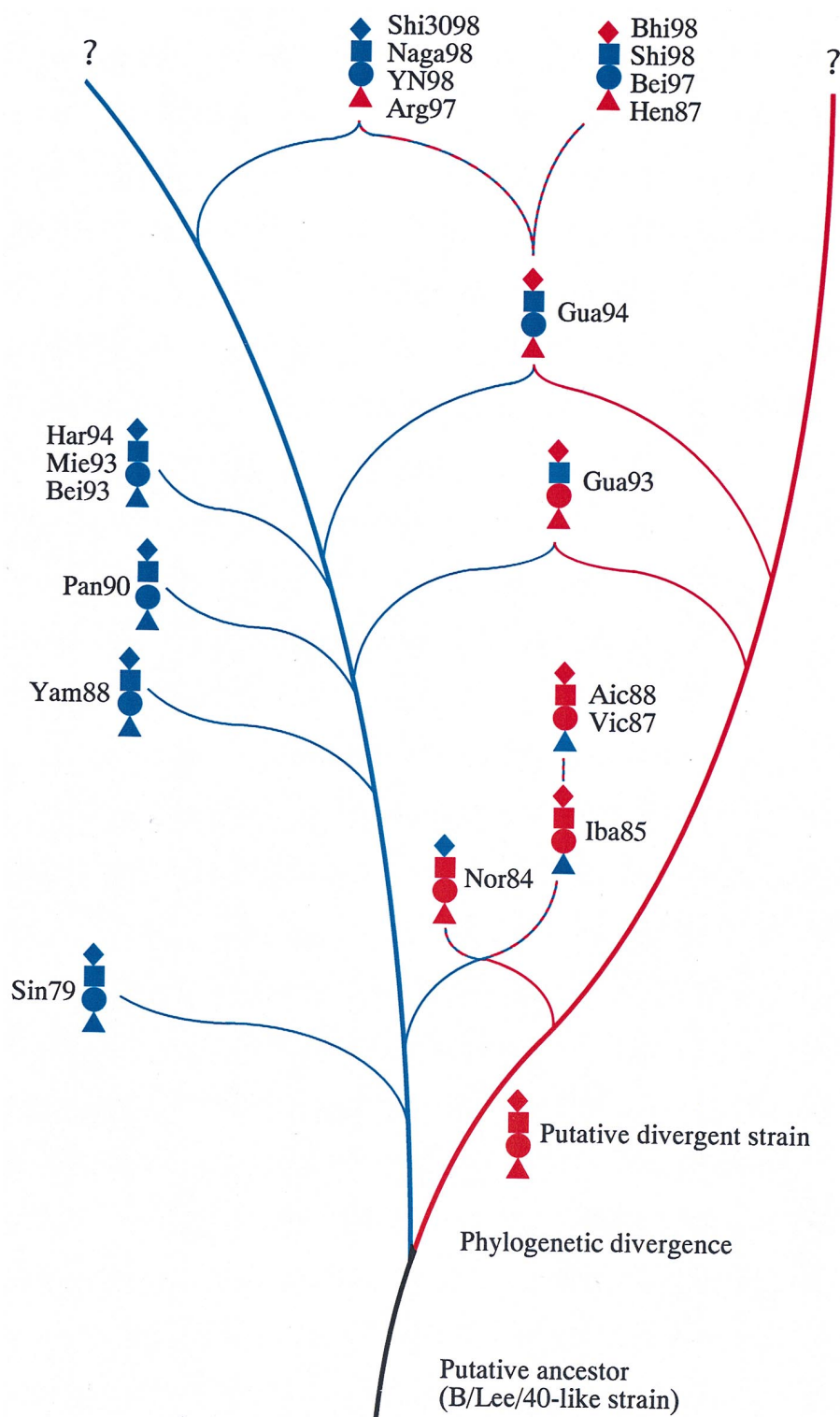
FIG. 5. Proposed pattern for evolution and gene reassortment among influenza B viruses. Genome constellations of influenza B viruses are represented as follows. The HA genes of lineages II and III are represented by blue diamonds and red diamonds, respectively; NP genes of lineages of II and III are represented by blue squares and red squares, respectively; M genes of lineages II and III are represented by blue circles and red circles, respectively; the corresponding lineages of NS genes (lineages III and IV) are symbolized by blue triangles and red triangles, respectively.

late in later viruses. This pattern was observed in lineage III of the evolutionary tree of the M gene as well as lineage IV of the NS gene. As shown in Fig. 5, genes of NS lineage IV first appeared in 1984, after which time genes of this lineage were not observed in circulating viruses for a period of 9 years. However, as represented by the NS gene of Gua93, genes of this lineage reemerged in 1993 and subsequently circulated in viruses from 1994, 1997, and 1998. Conversely, genes of lineage III were found in most isolates from the 1980s and early 1990s but were not observed in viruses isolated after 1994 (Har94) examined in this report. In a similar fashion, M genes of lineage III reappeared in 1993 (Gua93) after a 5-year hiatus (Aic88) but were not observed in later isolates. Sequestration and reemergence of genetic lineages through genetic exchange among circulating B viruses make prediction of genomic constellations of future influenza B viruses very difficult, if not impractical. Nonetheless, the question as to how these lineages are being sequestered is raised. Possible suggestions include subclinical infections by influenza B viruses which would not be detected, as well as circulation of influenza B viruses in an as-yet-undetermined host. However, because sequence data for the internal genes of these viruses is quite limited, the possibility that viruses containing genes of sequestered lineages have simply not been analyzed must also be considered.

Essentially identical topologies of the phylogenetic trees of the M and NP genes are noteworthy as they are typical of genes which are physically linked, suggesting a possible association between the M and NP proteins. The deduced M1 proteins of genes of both lineages, however, showed no conserved amino acid differences, while the BM2 protein unexpectedly demonstrated exceptionally high amino acid variability which was actually higher than that of the HA protein. Therefore, it was not only interesting to observe that the most variable (BM2) and least variable (M1) proteins were encoded on the same genome segment but that similar evolutionary patterns of the M and NP gene segments would more likely be the result of a codependent association between the NP and BM2 proteins than between the NP and M1 proteins. Further investigation into the role of the BM2 protein regarding its possible association with the NP protein would be warranted.

The evolutionary rates of the HA, NP, M, and NS genes of influenza B viruses were determined to be generally lower than those of their influenza A virus counterparts. Despite this fact, influenza B viruses were found to employ a series of evolutionary mechanisms which contributed to the variability of these viruses, including (i) genetic divergence of genes coding for the surface HA protein and internal proteins into two distinct major lineages, (ii) cocirculation of multiple lineages for extended periods of time, (iii) frequent reassortment among circulating viruses giving rise to new variants with distinct genome constellations, (iv) sequestration and reemergence of gene lineages, and (v) systematic deletion and insertion of nucleotides in the HA gene. Thus, despite the lower evolutionary rates of influenza B viruses, it is proposed that, through employment of these mechanisms, new variants of influenza B viruses with distinct genome constellations are generated, resulting in seasonal epidemic episodes. Furthermore, such complex phylogenetic patterns may be indicative of a symbiotic relationship among divergent viruses where timely exchange of gene segments provides the variability necessary for these viruses to continue circulating. This report demonstrates the fundamental differences between the evolutionary mechanisms of human influenza A and B viruses. Through ongoing global surveillance of influenza viruses, and genetic analysis of the genes coding not only for the surface HA protein but also for the internal proteins of these viruses, we can better understand the evolutionary mechanisms of these viruses. Further analysis of the remaining genes of influenza B viruses (NA, PA, PB1, and PB2 genes) would be worthwhile to better characterize the phylogenetic and reassortment patterns of these viruses.

REFERENCES

1. Betakova, T., M. V. Nermut, and A. J. Hay. 1996. The NB protein is an integral component of the membrane of influenza B virus. J. Gen. Virol. 77:2689–2694.
2. Brassard, D. L., G. P. Leser, and R. A. Lamb. 1996. Influenza B virus NB glycoprotein is a component of the virion. Virology 220:350–360.
3. Briedis, D. J., and R. A. Lamb. 1982. Influenza B virus genome: sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS1 and NS2 proteins. J. Virol. 42:186–193.
4. Briedis, D. J., R. A. Lamb, and P. W. Choppin. 1982. Sequence of RNA segment 7 of the influenza B virus genome: partial amino acid homology between the membrane proteins (M1) of influenza A and B viruses and conservation of a second open reading frame. Virology 116:581–588.
5. Briedis, D. J., and M. Tobin. 1984. Influenza B virus genome: complete nucleotide sequence of the influenza B/Lee/40 virus genome RNA segment 5 encoding the nucleoprotein and comparison with the B/Singapore/222/79 nucleoprotein. Virology 133:448–455.
6. Buonagurio, D. H., S. Nakada, W. M. Fitch, and P. Palese. 1986. Epidemiology of influenza C virus in man: multiple evolutionary lineages and low rate of change. Virology 153:12–21.
7. DeBorde, D. C., A. M. Donabedian, M. L. Herlocher, C. W. Naeve, and H. F. Maassab. 1988. Sequence comparison of wild-type and cold-adapted B/Ann Arbor/1/66 influenza virus genes. Virology 163:429–443.
8. Ellis, J. S., P. Chakraverty, and J. P. Clewley. 1995. Genetic and antigenic variation in the haemagglutinin of recently circulating human influenza A (H3N2) viruses in the United Kingdom. Arch. Virol. 140:1889–1904.
9. Fitch, W. M., R. M. Bush, C. A. Bender, and N. J. Cox. 1997. Long term trends in the evolution of H(3) HA1 human influenza type A. Proc. Natl. Acad. Sci. USA 94:7712–7718.
10. Fitch, W. M., J. M. Leiter, X. Q. Li, and P. Palese. 1991. Positive Darwinian evolution in human influenza A viruses. Proc. Natl. Acad. Sci. USA 88:4270–4274.
11. Geraci, J. R., D. J. St. Aubin, I. K. Barker, R. G. Webster, V. S. Hinshaw, W. J. Bean, H. L. Ruhnke, J. H. Prescott, G. Early, A. S. Baker, S. Madoff, and R. T. Schooley. 1982. Mass mortality of harbor seals: pneumonia associated with influenza A virus. Science 215:1129–1131.
12. Gojobori, T., K. Ishii, and M. Nei. 1982. Estimation of average number of nucleotide substitutions when the rate of substitution varies with nucleotide. J. Mol. Evol. 18:414–423.
13. Gojobori, T., E. N. Moriyama, and M. Kimura. 1990. Molecular clock of viral evolution, and the neutral theory. Proc. Natl. Acad. Sci. USA 87:10015–10018.
14. Hiebert, S. W., M. A. Williams, and R. A. Lamb. 1986. Nucleotide sequence of RNA segment 7 of influenza B/Singapore/222/79: maintenance of a second large open reading frame. Virology 155:747–751.
15. Hinshaw, V. S., W. J. Bean, J. Geraci, P. Fiorelli, G. Early, and R. G. Webster. 1986. Characterization of two influenza A viruses from a pilot whale. J. Virol. 58:655–656.
16. Horvath, C. M., M. A. Williams, and R. A. Lamb. 1990. Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide. EMBO J. 9:2639–2647.
17. Inglis, S. C., and C. M. Brown. 1981. Spliced and unspliced RNAs encoded by virion RNA segment 7 of influenza virus. Nucleic Acids Res. 9:2727–2740.
18. Ito, T., O. T. Gorman, Y. Kawaoka, W. J. Bean, and R. G. Webster. 1991. Evolutionary analysis of the influenza A virus M gene with comparison of the M1 and M2 proteins. J. Virol. 65:5491–5498.
19. Jambrina, E., J. Barcena, O. Uez, and A. Portela. 1997. The three subunits of the polymerase and the nucleoprotein of influenza B virus are the minimum set of viral proteins required for expression of a model RNA template. Virology 235:209–217.
20. Kanegae, Y., S. Sugita, A. Endo, M. Ishida, S. Senya, K. Osaka, K. Nerome, and A. Oya. 1990. Evolutionary pattern of the haemagglutinin gene of influenza B viruses isolated in Japan: cocirculating lineages in the same epidemic season. J. Virol. 64:2860–2865.
21. Kinnunen, L., N. Ikonen, T. Poyry, and R. Pyhala. 1992. Evolution of influenza B/Victoria/2/87-like viruses: occurrence of a genetically conserved virus

under conditions of low epidemic activity. J. Gen. Virol. **73:**733–736.

22. **Klingborn, B., L. England, R. Rott, N. Juntti, and G. Rockborn.** 1985. An avian influenza A virus killing a mammalian species—the mink. Arch. Virol. **86:**347–351.

23. **Krystal, M., J. F. Young, P. Palese, I. A. Wilson, J. J. Skehel, and D. C. Wiley.** 1983. Sequential mutations in hemagglutinins of influenza B virus isolates: definition of antigenic domains. Proc. Natl. Acad. Sci. USA **80:**4527–4531.

24. **Lamb, R. A., S. L. Zebedee, and C. D. Richardson.** 1985. Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. Cell **40:**627–633.

25. **Lindstrom, S., A. Endo, S. Sugita, M. Pecoraro, Y. Hiromoto, M. Kamada, T. Takahashi, and K. Nerome.** 1998. Phylogenetic analyses of the matrix and non-structural genes of equine influenza viruses. Arch. Virol. **143:**1585–1598.

26. **Lindstrom, S. E., Y. Hiromoto, R. Nerome, K. Omoe, S. Sugita, Y. Yamazaki, T. Takahashi, and K. Nerome.** 1998. Phylogenetic analysis of the entire genome of influenza A (H3N2) viruses from Japan: evidence for genetic reassortment of the six internal genes. J. Virol. **72:**8021–8031.

27. **Lindstrom, S. E., S. Sugita, A. Endo, M. Ishida, P. Huang, S. H. Xi, and K. Nerome.** 1996. Evolutionary characterization of H3N2 influenza viruses: novel changes in the receptor binding domain. Arch. Virol. **141:**1349–1355.

28. **Londo, D. R., A. R. Davis, and D. P. Nayak.** 1983. Complete nucleotide sequence of the nucleoprotein gene of influenza B virus. J. Virol. **47:**642–648.

29. **Ludwig, S., U. Schultz, J. Mandler, W. M. Fitch, and C. Scholtissek.** 1991. Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses. Virology **183:**566–577.

30. **Murphy, B. R., and R. Webster.** 1996. Orthomyxoviruses, p. 1397–1446. *In* B. N. Fields, D. M. Knipe, P. M. Howley, R. M. Chanock, J. L. Melnick, T. P. Monath, B. Roizman, and S. E. Staus (ed.), Fields virology, 3rd ed. Lippincott-Raven, Philadelphia, Pa.

31. **Nei, M., and T. Gogobori.** 1986. Simple methods for estimating the number of synonymous nucleotide substitutions. Mol. Biol. Evol. **34:**418–426.

32. **Nerome, R., Y. Hiromoto, S. Sugita, N. Tanabe, M. Ishida, M. Matsumoto, S. E. Lindstrom, T. Takahashi, and K. Nerome.** 1998. Evolutionary characteristics of influenza B virus since its first isolation in 1940: dynamic circulation of deletion and insertion mechanism. Arch. Virol. **143:**1569–1583.

33. **Norton, G. P., T. Tanaka, K. Tobita, S. Nakada, D. A. Buonagurio, D. Greenspan, M. Krystal, and P. Palese.** 1987. Infectious influenza A and B virus variants with long carboxyl terminal deletions in the NS1 polypeptides. Virology **156:**204–213.

34. **Peng, G., S. Hongo, H. Kimura, Y. Muraki, K. Sugawara, F. Kitame, Y.** Numazaki, H. Suzuki, and K. Nakamura. 1996. Frequent occurrence of genetic reassortment between influenza C virus strains in nature. J. Gen. Virol. **77:**1489–1492.

35. **Robbins, P. A., P. A. Rota, and S. Z. Shapiro.** 1997. A broad cytotoxic T lymphocyte response to influenza type B virus presented by multiple HLA molecules. Int. Immunol. **9:**815–823.

36. **Rohm, C., N. Zhou, J. Suss, J. Mackenzie, and R. G. Webster.** 1996. Characterization of a novel influenza hemagglutinin, H15: criteria for determination of influenza A subtypes. Virology **217:**508–516.

37. **Rota, P. A.** 1989. Sequence of a cDNA clone of the nucleoprotein gene of influenza B/Ann Arbor/1/86. Nucleic Acids Res. **17:**5395.

38. **Rota, P. A., M. L. Hemphill, T. Whistler, H. L. Regnery, and A. P. Kendal.** 1992. Antigenic and genetic characterization of the haemagglutinins of recent cocirculating strains of influenza B virus. J. Gen. Virol. **73:**2737–2742.

39. **Rota, P. A., T. R. Wallis, M. W. Harmon, J. S. Rota, A. P. Kendal, and K. Nerome.** 1990. Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. Virology **175:**59–68.

40. **Saitou, M., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.

41. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **14:**5463–5467.

42. **Shimbo, K., D. L. Brassard, R. A. Lamb, and L. H. Pinto.** 1995. Viral and cellular small integral membrane proteins can modify ion channels endogenous to Xenopus oocytes. Biophys. J. **69:**1819–1829.

43. **Shu, L. L., W. J. Bean, and R. G. Webster.** 1993. Analysis of the evolution and variation of the human influenza A virus nucleoprotein gene from 1933 to 1990. J. Virol. **67:**2723–2729.

44. **Sundstrom, N. A., L. S. Premkumar, A. Premkumar, G. Ewart, G. B. Cox, and P. W. Gage.** 1996. Ion channels formed by NB, an influenza B virus protein. J. Membr. Biol. **150:**127–132.

45. **Verhoeyen, M., L. van Rompuy, J. W. Min, D. Huylebroeck, and W. Fiers.** 1983. Complete nucleotide sequence of the influenza B/Singapore/222/79 virus haemagglutinin gene and comparison with the B/Lee/40 haemagglutinin. J. Gen. Virol. **73:**2737–2742.

46. **Yamashita, M., M. Krystal, W. M. Fitch, and P. Palese.** 1988. Influenza B virus evolution: co-circulating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. Virology **163:**112–122.

47. **Zou, S.** 1997. A practical approach to genetic screening for influenza virus variants. J. Clin. Microbiol. **35:**2623–2627.